This was quite the challenging project, and my final results are far from ideal. Here were the challenges I faced:

**Data Preprocessing**

I was unclear as to how to preprocess the data as I realize if I stripped the data of all of its qualities, the model wouldn’t receive data that still retained the stylistic qualities of a standard Shakespeare text. For example, stemming would cut down words to its core component and in this case, I assumed it was not the way to go. Therefore, I skipped stemming and the only data preprocessing I did was clean strip out whitespaces, but I kept the new lines.

**Tokenizing using BPE**

Wasn’t sure what vocab size to use here because the main concern for me was computing time as the GPU in my system is weak and Colab only gives out a limited time of GPU usage (we need GPU handouts from MMA!). So, I stuck with 1000 vocab size, but I would like to see how much better my results would be if my vocab size was something like 5000 (which is what GPT recommended).

**Creating the RNN architecture**

Had lots of trouble here trying to find the best parameters here while also not setting numbers that would take forever for the model to train. In the end, I chose batch size of 64 (which is too high but had reasonable performance time), 1 epoch, learning rate of 0.01, embedding dimensions 128 and hidden dim 256. I firmly believe that if these numbers were higher my models would perform better.

**Results**

In terms of perplexity, it seems LSTM had the best perplexity but in terms of text generation GRU appeared to have performed the best. The sentences didn’t really make sense but at least there were words that made sense. This was the result from the GRU model:

*To be , or not to be . unto myself nor R fi ble nor nor take i nor MEN say there S boy -- help being thr friends away Shall nor his P R him from up ROMEO nor nor nor pri nor know pri er nor wife un st sub Eng ter ans , tem hath lie n this ted ps right ent never for sent pat gracious else ment I nor WARWICK nor sa love un now LEONTES let down . ange but before more house s wr ion Why s I this si y s , pl all y ate lack st ion*

**Beam search results were not good**

Need to look more into why my beam search did not perform well.